SPATIOTEMPORAL BAYESIAN CELL POPULATION TRACKING AND ANALYSIS WITH LINEAGE CONSTRUCTION

Luke M. A. Beaumont[1,3]*, James Wakefield*[2,3] *and J. Alison Noble*[1]

[1]Wolfson Medical Vision Laboratory, Dept. of Engineering Science, University of Oxford, UK
[2]Dept. of Zoology, University of Oxford, UK
[3]Life Sciences Interface: Doctoral Training Centre, University of Oxford, UK

ABSTRACT

Tracking of cell populations *in vitro* in time lapse microscopy images enables automatic high throughput spatiotemporal measurements of a range of cell cycle mechanics and dynamics. Both in clinical and academic environments, large scale cellular data analysis using such methods stands to facilitate a paradigm shift in approaches to understanding cell biology. In this paper, we present a novel approach to cell population tracking and segmentation. We employ the CONDENSATION algorithm in tandem with Fast Levels Sets and Exclusion Zones for robust tracking and pixel-accurate segmentation. The algorithm feeds its output to a lineage filter. The complete approach is validated in terms of its ability to track and identify nuclei, and by its success in detecting abnormalities in the length of mitosis.

Index Terms— Spatiotemporal, Bayesian, Cells, Analysis.

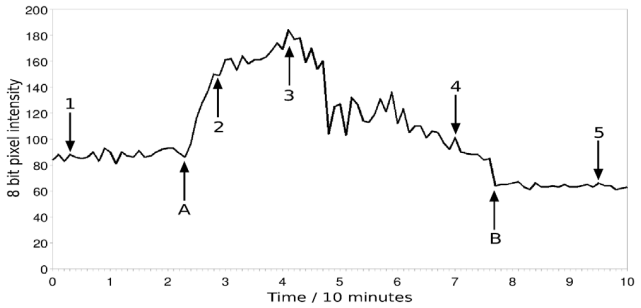
1. INTRODUCTION

Automatic cell tracking and segmentation are a pre-requisite to automatic cell analysis in image sequences. Manual methods can never sustainably analyse the volume of data required by high-throughput methods. Commercially available microscopy image processing systems mainly employ simple algorithms, such as thresholding techniques that are too simplistic to provide robust tracking. Hence, we are interested in more sophisticated approaches based on a knowledge of the imaging process and the underlying biological mechanisms. Our work is motivated by cell biology and a need to measure cell cycle parameters. For instance, the spindle assembly checkpoint (SAC) is a component of the cell cycle that prevents a cell from physically dividing its nuclear material until all chromosomes are in a position such that they will faithfully divide. Mutations in SAC are frequent in all cancers. If SAC is inactive, the cell will proceed through mitosis more quickly than usual. With SAC active, problems with the correct assembly of cell division components will maintain an active checkpoint, and mitosis will continue to persist. Eventually, usually within 24 hours, the cell will undergo programmed cell death or *apoptosis* without any detriment to the surrounding cells. SAC is therefore an obvious target for drug treatment to arrest cancerous cell development. In this paper we consider the intensity profile of a nucleus through the cell cycle (Fig.1) as a means to derive the length of mitosis automatically, and therefore detect abnormal cells.

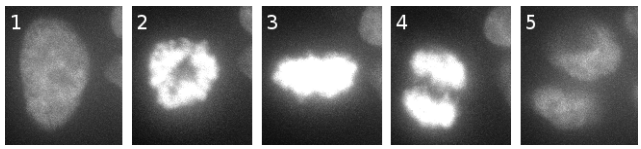
Automatic microscopy analysis is an emerging area of biomedical image analysis. For instance, Li et al.[1] describe a track-linking based approach to spatiotemporal tracking and lineage construction in cells. The approach compares the outputs of a level set segmentation and an intensity histogram based segmentation to identify candidate cells in an individual image. These are then connected between frames using an Interacting Multiple Model (IMM) filter. Tracking accuracies between 83.8% and 92.5% are reported on MG-63 osteosarcoma (cancer) cells over 10 hours and on aminon epithelial (AE) stem cells over 42.5 hours. Chen et al.[2] employ a nuclei-fragment merging technique based on prior knowledge of nuclei shape and size to reduce over-segmentation in nuclei. A K-nearest neighbour classifier is used to determine cell phase. Tracking accuracies of 89.3% to 94.3% are reported over 96 frames over 24 hours. Our approach differs from the above in the use of a hybrid CONDENSATION/Level Sets algorithm and embedding knowledge of biological mechanism models and image formation in the solution.

2. UNDERSTANDING THE IMAGES

Histone GFP Telomerase HeLa cells are an immortal cell line bred specifically for use in cancer research [3]. Histones are protein compounds that form a part of the chromosomes, and as such in histone GFP HeLa cells visualisation of the nucleus throughout the cell cycle is possible. Although now heavily mutated from their origins as human cells, for imaging purposes HeLa cells provide an accurate model. Moreover, the availability of HeLa cell lines stably expressing fluorescently tagged cellular proteins allows specific processes within the cell to be monitored. Histones are essential, evolutionarily conserved proteins that bind to DNA, forming chromosomes, and as such, histone GFP HeLa cells allow the visualisation of the nucleus throughout the cell cycle.



(a)



(b)

Fig. 1. *(a) Intensity profile for a typical nucleus undergoing mitosis. (b) The corresponding confocal microscope images.*

Intensity profile. Images of a typical HeLa cell and its associated intensity profile for one mitotic division can be seen in Fig.1. For the majority of the cell cycle, the cell replicates DNA and slowly increases in size as it prepares to undergo mitosis and be divided in two (image 1 and 5). During this period, although the observed size of the nucleus increases, the density of the chromosomes and therefore the observed average intensity of the nucleus remains approximately constant.

Throughout mitosis (between A and B in Fig.1(a)), the chromosomes condense, and so the concentration of GFP increases resulting in a pronounced increase in intensity. Initial condensation is rapid, with clearly observable changes happening within the order of five minutes of the onset of mitosis (shown between A and 2). Intensity peaks just before the physical separation of nuclear material, known as *Anaphase* (3). During Anaphase itself (4) there is a pronounced reduction in intensity followed by a fade to a stable non-mitotic intensity (5) over the period of approximately thirty minutes. Mitosis usually lasts about an hour.

Morphological profile. Outside of mitosis, the cell and nucleus slowly increase in volume in preparation for division. The nucleus appears as an approximate circle which gradually increases in area. In contrast, the events of mitosis are more pronounced, key events in mitosis are shown in Fig.1. As the cell prepares to divide, condensed chromosomes are aligned along an axis, they appear as a bright ellipse on microscope images. At Anaphase, the chromosomes are pulled apart and so the ellipse splits along its major axis. The divided chromosomes then decondense as they form part of the nucleus of their new cell. The bright ellipses fade as they adopt a more circular shape. If we can find A and B we can automatically derive the length of mitosis.

3. ALGORITHM

The new algorithm is outlined in Fig.2, with the key tracking steps highlighted.

3.1. Local adaptive thresholding

We employ Local Adaptive Thresholding (LAT) [4] to mitigate the effect of regional variations in intensity caused by nuclei at varying focal planes. LAT calculates an individual threshold value for a pixel based on the intensity profile of a window around the pixel. LAT requires a single parameter based on a global threshold value. The Otsu algorithm [5], based on a two Gaussian intensity profile was found to provide a good global threshold. Images are also median filtered to remove the resulting binarised small unconnected pixels.

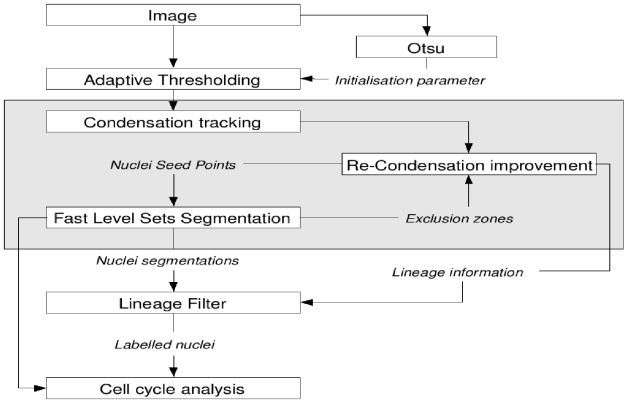


Fig. 2. *Algorithm overview*

3.2. CONDENSATION Tracking

Cell tracking is achieved using a variant of the CONDENSATION algorithm [6]. The algorithm is based on random sampling, where the chance of selecting a particular sample is proportional to how well the sample's observed data fits a model of the object being searched for.

We model the shape of a nucleus throughout the cell cycle as an ellipse (section 2). The major/minor axes of the model are bounded by the minimum/maximum size of a nucleus at any point in the cell cycle. The orientation can take any value.

A sample's position, size and orientation are initially chosen from a uniform distribution. Observation measurements are taken by edge detectors, evenly spaced along the length of the ellipse. Each of these searches for edges in the thresholded image, and returns a score based on the location of a detected edge. A sample's score of how well its observed data fits the model is the sum of these values.

Accordingly, samples located on the image at a nucleus and with the appropriate size and orientation will score highly.

Samples away from nuclei will return a low score. A sample's chances of being selected for the next frame are directly proportional to the score in the current frame.

To predict movement between frames, a selected sample's position, orientation and size are adjusted in accordance with an expected movement model, based on a Gaussian distribution.

Outside of mitosis, HeLa cells move very little, and the movement model can be assumed Gaussian with a small σ. More dramatic movement occurs during mitosis when the nucleus changes its morphology and divides. During this period, a larger σ is required to follow the exaggerated movement. The transition point between these two states is detected from the change in the intensity of the nucleus.

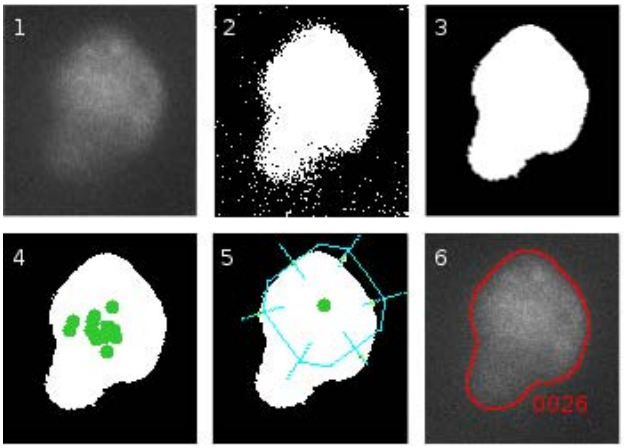


Fig. 3. *Stages of image processing on an individual nucleus for a single frame. 1. Original image. 2. Adaptive thresholded. 3. Median filtered. 4. Condensation tracking according to posterior likelihood sampling. 5. The top scoring sample's spline. 6. The top scoring sample is used to initialise a Fast Level Set to segment the nucleus. Lineage construction uniquely identifies the nucleus between frames.*

3.3. Exclusion Zones and redistribution

With a completed iteration of CONDENSATION, we have an image with a high concentration of samples over the best fitting objects. To force all objects to be detected, we record high scoring areas and then redistribute non-local maxima score samples from these areas to other parts of the image. The highest scoring sample in the image is selected, and used to seed a Fast Level Sets segmentation. All the other samples that reside within the segmented region are removed and added to a list for redistribution. A proportion, γ, of the combined score of the removed samples is added to that of the remaining sample. In this way we minimise the effect on the probability density for the next image. For the redistribution of removed samples, the weight for the top scoring sample is set to zero, and as such the Exclusion Zone is established. This process is repeated for the next remaining highest scoring sample and so on. The result is an image with a number of nuclei indicated by one sample, a probability density with corresponding regions set to zero and a list of unused samples ready for redistribution.

Redistribution of samples can happen in two ways - resampling and uniform redistribution. The latter is uniformly randomly positioning the samples, as on the first initialisation image. This mechanism allows for the identification of newly appeared nuclei or of any nuclei previously lost. The former, resampling, involves sampling from the sample set from the previous image. With Exclusion Zones established, this forces the algorithm to select samples with a relatively smaller likelihood, which have now become more dominating. This mechanism ensures the algorithm keeps track of all nuclei, without cost to the weight or number of samples apportioned to the best fitting samples. A quality threshold prevents noise elements being tracked.

3.4. Fast level sets

Fast Levels Sets [7] is used to produce an accurate segmentation of a nucleus from a single starting point inside the nucleus image. An example of such a segmentation can be seen in Fig.3. The segmentation represents the entire and accurate region which an individual sample represents. It is this region which forms an *Exclusion Zone* to allow multiple nuclei to be tracked, and is used for subsequent analysis.

3.5. Lineage construction

CONDENSATION is a generational algorithm that allows the construction of a temporal model of cell behaviour (lineage). The child of a sample on frame n seeks the same nucleus on frame $n+1$ and similarly, the grandchildren will seek the same nucleus on from $n+2$. Lineage information is passed between frames innately by the algorithm. Occasionally, however, nuclei may partially or totally occlude one another or may move in and out of the viewing area. This presents difficulties for the preservation of lineage as when two groups of samples operate in the same area one group will be removed by an *Exclusion Zone*. If the objects subsequently move out of occlusion, their lineage information is lost. To account for these situations, the method of Li et al. [1] is used to connect lost tracks.

3.6. Cellular analysis

For individual nuclei, identifying the length of mitosis equates to finding the period between A and B on Fig.1(a). This region is detected by local intensity gradient analysis. The σ for cells undergoing normal mitosis is very small. We identify abnormal mitotic events as those that differ from the mean by more than 3σ.

4. RESULTS

Datasets. We analysed the performance of the approach on four image sequences of histone-2B-GFP HeLa cells taken on a Deltavision RT laser scanning microscope using a 40x objective. All images were taken at a temporal resolution of one minute, with three image sets collected over 200 minutes and one over 150 minutes. The images were recorded at a spatial resolution of 896 x 896 pixels.

Parameter values. The approach has parameter values associated with the ellipse based model of a nucleus and its predicted movement, with the function of the CONDENSATION algorithm, and with the function of the Fast Level Sets algorithm. The parameter values were determined by experimentation for best performance in each case.

The length of an ellipse minor axis was bounded between 10 and 40 pixels and the major axis between 30 and 40 pixels. Initialisation prior mean and σ values were 35 and 1 pixels for the major axis and 35 and 5 pixels for the minor axis respectively. For rotation the prior mean value was π from the vertical with a σ of $\frac{\pi}{2}$. For the movement prediction model, the σ for position, rotation minor and major axes were 6.5 pixels, $\frac{\pi}{15}$, 1 pixels and 0.1 pixels respectively. There were 500 samples in the CONDENSATION sample set. The LAT window size was square and set to 100 x 100 pixels. The Exclusion Zone γ value was 0.5.

Validation. The tracking program was set to label individual nuclei with a unique number on every frame in each image sequence. A human operator manually tracked each nucleus and checked the identifier on every frame. Tracking of a nucleus was only considered valid if the *same* nucleus was tracked throughout the entire period it was visible on the image sequence. Intensity profiles for cells that went through nuclear division were manually checked for entry and exit points into mitosis. Table 1 summarizes the results.

Table 1. Tracking accuracy results

Sequence	Lineage validity	Occluded nuclei removed
A	26/28 (92.9%)	26/27 (96.3%)
B	32/35 (91.4%)	32/32 (100.0%)
C	32/38 (84.2%)	32/33 (96.9%)
D	34/39 (87.2%)	34/34 (100.0%)
All	*124/140 (88.6%)*	*124/126 (98.4%)*

The vast majority of lineage tracking failure occurred when a nucleus became totally or partially occluded for a sustained period - longer than 4 frames. This would cause all lineage information to be lost for the occluded nucleus. 98.4% of nuclei were tracked accurately in the absence of extended occlusion.

Table 2 summarises the investigation on mitosis length estimation. 22 normal and 3 abnormal mitotic events were identified. The distribution of results for normal nuclei showed a small σ of 2.39 minutes. Abnormal events were identified as lying outside 3σ.

Table 2. Mitosis lengths

Seq.	Normal events	Avg. len. *(Mins.)*	Abnormal events	Len. *(Mins.)*
A	5	53.8	0	
B	10	58.2	1	85*
C	4	56.5	1	116
D	3	55.0	1	94*
All	*22*	*56.5 (σ = 2.39)*	*3*	

** nuclei were still in mitosis at the end of the image sequence*

5. CONCLUSIONS AND FURTHER WORK

We have developed and validated a new algorithm for tracking and tracing the lineage of dense populations of nuclei in confocal microscope images of HeLa cells. The combination of lineage information and pixel accurate nuclei segmentations were used to automatically identify abnormalities in the length of the cell cycle. Our results match the best claimed tracking performance in the field [1, 2], though differences in the nature of images used for testing make direct comparison difficult. Future work includes further development of the algorithm to handle extended occlusion, measurement of other key cell parameters, and deployment of the method for large scale cell characterisation studies.

We are indebted to Jennifer Winter and the McAinsh Laboratory at the Mari Curi Research Institute for their generous provision of the images used in this research.

6. REFERENCES

[1] K.Li et al., "Cell population tracking and lineage construction with spatiotemporal context," in *MICCAI*, 2007, pp. 295–302.

[2] X.Chen et al., "Automated segmentation, classification and tracking of nuclei in time-lapse microscopy," *TBME*, vol. 53, no. 4, pp. 762–6, 2006.

[3] J.Masters, "Hela cells 50 years on: the good, the bad and the ugly," *Nat Rev Cancer.*, vol. 2, no. 4, pp. 315–9, 2002.

[4] E. Davis, *Machine Vision:Theory, Algorithms, Practicalities*, Academic Press, 1990.

[5] N. Otsu, "A threshold selection method from gray-level histograms," in *TSMC*, 1979, vol. 9, pp. 62–6.

[6] A. Blake and M. Isard, "The CONDENSATION algorithm - conditional density propagation and applications to visual tracking," in *NIPS*, 1997, vol. 9, p. 361.

[7] Y.Shi et al., "A Fast implementation of the Level Set method without solving PDEs," in *ICASSP*, 2005, vol. 2, pp. 97–100.